HOW MANY CATEGORIES ARE IN VERY HIGH RESOLUTION SAR IMAGES?

Corneliu Octavian Dumitru, Mihai Datcu

Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Wessling, Germany

ABSTRACT

In this paper, we propose to identify the number of categories that can be retrieved from a very high resolution SAR data. The evaluation is done on TerraSAR-X high resolution Spotlight data and the retrieved categories are semi-automatically annotated using as feature vector the Gabor filters; as a classifier the Support Vector Machine, and for ranking the suggested images the relevance feedback. The visualization of the tool was enhanced compared with our previous implementation in order to support the users in his/her approach to search the patches of interest in a large repository. Our dataset consist in 43 scenes that cover as much as possible all the regions over the world. A total of 352 categories are identified that contain urban and non-urban categories.

Index Terms—categories, Gabor filters, patch, semantic annotation, high resolution TerraSAR-X.

1. INTRODUCTION

Increasing the number of very high resolution Earth Observation satellites has a significant impact on the information content in images.

Until now, for satellite images (e.g., TerraSAR-X) the number of categories that can be identified in each image is very limited [1] and these categories are not semantically annotated contrary with multimedia where this stage is very well established and the identified categories are already annotated [2] ÷ [4]. For cartography [5], [6] some generic categories with their semantic annotation are defined for a limited number of areas in Europe and these categories are not matching with the case of very high resolution data.

In the last years the number of categories identified and annotated were limited and manually identified (e.g., in rapid mapping processing are used in average 10 categories). In our first demarche, we start with two scenes covering the Venice and Toulouse and from these two scenes about 30 categories were retrieved with an accuracy of precision / recall of 82.27% and 35.62% [7]. Later, we extended our dataset to 39 scenes and 320 categories were retrieved during the evaluation [8], [9].

In this paper, we progress compared to what was done before by adding new semantically annotated scenes to our dataset. We also presented the evolution of the number of categories identified from an enhanced dataset of 43 new TerraSAR-X high resolution images available for the second phase of our project [9] compared with the number of categories identified in our previous phase of the project where only 39 scenes were considered for annotation.

The paper structure is the following. Section 2 briefly describes the evolution from phase I to phase II of the project where the number of scenes was increased from 39 to 82 scenes. Section 3 explains the methodology used to semantically annotate the categories extracted from very high resolution TerraSAR-X data. Section 4 presents the characteristics of the enlarged TerraSAR-X dataset. Section 5 provides two examples that are obtained after the semi-automated methodology was applied to the enlarged dataset. The paper ends with conclusions in Section 6.

2. EVOLUTION FROM PHASE I TO PHASE II

In Table 1 we show the two phases of the "Ontologies for the VO for TerraSAR-X data" deliverable of our project [9] by presenting the number of scenes annotated and the number of retrieved categories in each phase.

During Phase I, 39 scenes were downloaded from the TerraSAR-X archives via the EOWEB portal and processed using a Support Vector Machine with Relevance Feedback (SVM-RF). The scenes were tiled in patches and all patches were annotated.

In the second phase, another 43 scenes were downloaded and about 36,342 patches were grouped into 352 categories using the SVM-RF tool with an enhanced visualisation (see the next section).

Table 1: Semantic annotation for Phase I and II.

PHASE	SCENES / PATCHES	SEMANTIC CATEGORIES
I	39 SCENES 40,307 PATCHES	336 CATEGORIES
I+II	82 SCENES 76,649 PATCHES	688 CATEGORIES

In Figure 1 is presented the distribution of the dataset that covers different areas over the world for both phases and in Figure 2 the location of these scenes is marked with red

colour (inside of a diamond). In the case when many scenes are available for the same country, these locations are marked on the map with only one red diamond.
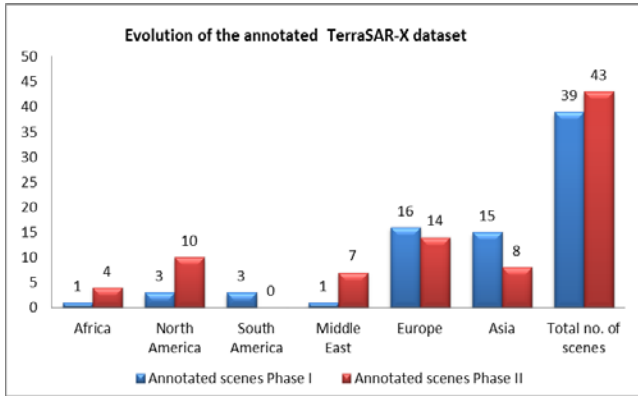

Figure 1: Evolution of the number of semantic annotated scenes from Phase I to Phase II.

3. METHODOLOGY

The steps of annotation methodology were presented in [9] and also in [10]. Here in the next paragraphs, we present an enhanced methodology in 7 steps.

In order to annotate the enhanced dataset the following steps shall be applied to each product:

1) Group the scenes (products) in collections. Collections means that one, two or more scenes are grouped together based on some criteria (e.g., geographically, architectural, etc.).

2) Select the optimal parameters of the data. A study was conducted [11], [12], [7] in order to find the best resolution, pixel spacing, optimal patch size, incidence angle, and orbit direction that corresponds to the basic detected TerraSAR-X products. Based on this study we selected: MGD products with radiometrically enhanced radiometry, high resolution Spotlight mode with single polarization acquisitions, the patch size of 160 by 160 pixels covering an area of 200m x 200m on ground, and the incidence angle close to the upper bound of the sensor range with an ascending or descending orbit looking.

3) Tile the MGD product-image into patches of 160 x 160 pixels (taking into account the number of the reference columns and rows – parameters that are available in the TerraSAR-X XML file [13]). These parameters from the metadata are needed in order to remove the black letter box from an image that appears in some images in the left or right part of the product-image depending on the orbit looking (ascending or descending).

4) Generate a quick-look (in "jpg" format without rescaling the data) of each patch and also a quick-look of the full image needed for visualization.

5) Compute the primitive features associated with each patch. Gabor filters are used as primitive features generators. A comparative study was presented in [7] for different PF algorithms that demonstrated that Gabor filter is the best algorithm. Our configuration of Gabor filters is 4 scale and 6 orientations [14] that gives us feature vectors of 48 components (computing the mean and variance of each scale and orientation).

6) Select a classifier in order to group the features into categories. For this issue we built a tool based on a Support Vector Machine (SVM) with a relevance feedback (RF). The SVM-RF tool supports users to search patches (the quick-look of these patches) of interest in a large repository having as a support on the right side a full image.
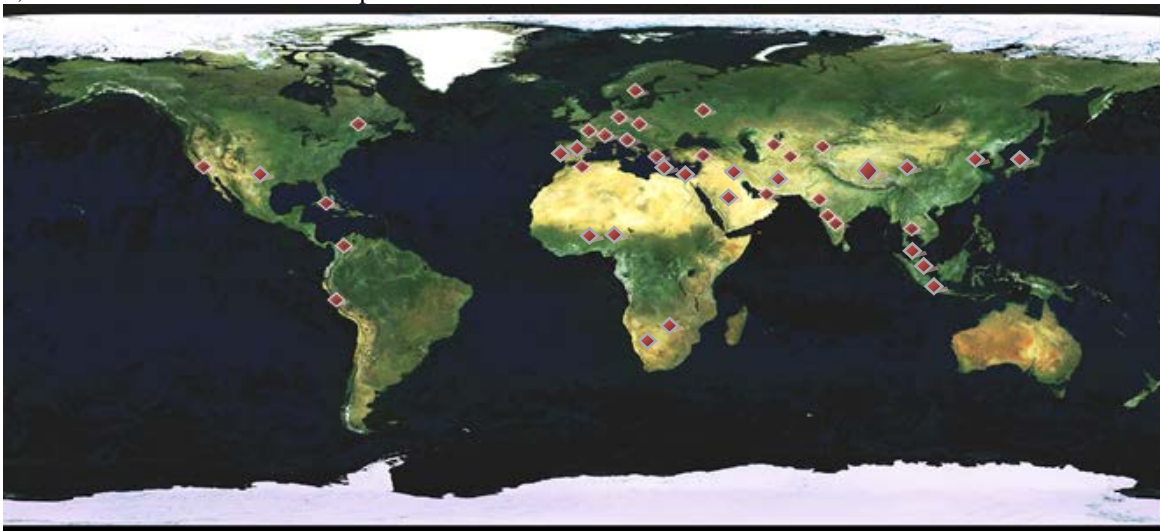

Figure 2: Location of each scene available in dataset (Phase I and II) marked on a map [15] with a red diamond. The scenes are distributed over the world.

The performances of this classifier, reported in the literature, are very good and the kernel has the capacity to perform highly accurate classification using a very limited number of examples.

The Graphical User Interface (GUI) of this tool allows Human-Machine Interaction (HMI) to rank the automatically suggested images which are expected to be grouped in the class of relevance.

In the proposed approach, patches are assigned to a single category based on the dominant content in the patch. The training and generating of categories stops after all patches have been grouped.

7) Annotate semantically each category using as visual support the ground truth of Google Earth and give an appropriate meaning to each category. After an appropriate label is found for each category the quick-looks that are belonging to this category are moved from the dataset into a folder bearing the name of the category. Once the generation of the category is finished, a new classification can start until all the patches from the dataset are annotated.

4. DATASET DESCRIPTION

Traditionally, image datasets are built by individual groups and are being used to solve specific problems. For different domains the available datasets may have a large variety in the number of retrieved categories.

Our *dataset* for the phase II contains 43 scenes over the world and the distribution of these scenes is presented (with red colour) in Figure 1.

The important characteristics of the TerraSAR-X dataset are the following: (1) the Multi-look Ground range Detected products, radiometrically enhanced, high resolution Spotlight mode; (2) the pixel spacing is equal to 1.25 meters, and the resolution about 2.9 meters; (3) the average of the scene sizes is 4200 x 5500 pixels and about 900 patches per scene are obtained after tiling (the patch size being 160 x 160 pixels); (4) the incidence angle between 30° and 50°; (5) compare with the previous dataset [8] in Phase II the orbit looking, for most of the scenes, is ascending but there are scenes with descending looking; and (6) the single polarization is HH or VV.

In order to understand the diversity of the categories that can exist in a remote sensing dataset, we present in Figure 3 a set of sub-images (called patches) tiled from the images being available. For very high resolution SAR images the diversity of the categories that can be retrieved from an image is higher than in the case of lower or medium resolution.

5. RESULTS

In this section, we present the taxonomies that are defined during the evaluation from Phase I [8] and used in Phase II
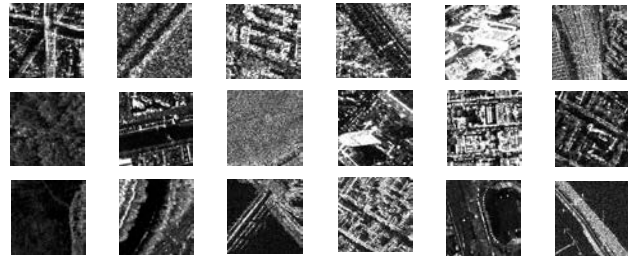

Figure 3: Diversity of categories that can be extracted from the available scenes.

of our project [9]. Based on the evaluation from previous phase, we arrived to a standard hierarchical annotation scheme. This scheme is a two-level annotation scheme: level 1 gives general information about the content of the patch (e.g., urban area, agriculture, transportation, forest, etc.) while level 2 details the general information from level 1 (e.g., urban area → high/medium/low density residential area, industrial area, ...; transportation → railway, port, road, bridge,...; etc.). Notice, however, in reality you can have patches whose semantics can refer to two or more detailed taxonomies from level 2.

The scenes presented in section 4 covers different areas over the world and some of them can be grouped together, for annotation, based on some defined criteria. Three criteria were introduced in [8], [9]:

- for the first criterion we considered only one scene for annotation – in order to have an idea about how many categories can be identified for each continent/region/country.

Twelve collections are generated based on the first criterion and this are covering the following areas: Bordeaux, Bydgoszcz, Jakarta, Kuala Lumpur, Madrid, Mueang Yala, Oslo, Porto, San Francisco, Singapore, Trento, and Van.

- for the second criterion two scenes are grouped together having different geographical location (e.g., country) – in order to see if is possible to have the same urban categories belonging to two different scenes.

Six collections are generated based on the second criterion and the covered areas are: Belgrade and Skopje, Chania and Thessaloniki, Baghdad and Bandar Imam Khomeini, Dhahran and Riyadh, Pyongyang and Suwon, Ashdod and Beirut.

- for the last criterion the grouping is made based on the geographical location of the scenes or based on similar architectural characteristics of the urban area – in order to help us to annotate large archives.

Four such collections are generated: two collections based on the geographical location of the scenes: North America (USA, Mexico, and Canada) and Africa (Nigeria, South Africa, Togo, and Zimbabwe) and other two collections based on the architectural characteristics: Germany

(Karlsruhe, Lindau, Mannheim, and Stuttgart) and Malaysia (Alor Setar and Seremban).

In next two figures are presented the categories retrieved and semantically annotated after the methodology presented in section 3 was applied to the selected dataset. In Figure 4, we considered the first criterion of grouping, while in Figure 5 the geographical location of the scenes was taking into a count as a part of the third criterion of grouping.
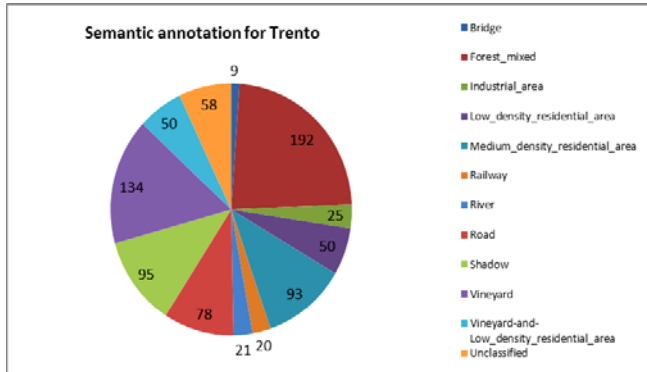

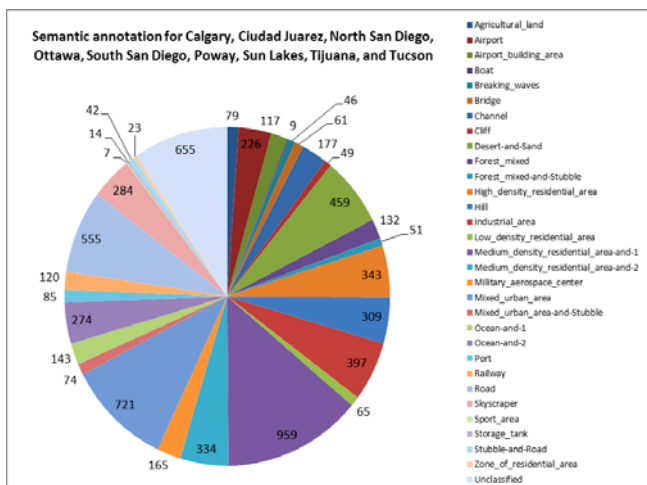**Figure 4: Number of patches per semantic category for the Trento image.**

Figure 5: Number of patches per semantic category retrieved after classification of nine scenes over North America.

6. CONCLUSION

This evaluation is very important for catalogue generation and for a complete annotation and understanding of the scenes.

With increasing the volume of data (number of scenes in the dataset) the manual interpretation of the data will be not possible and automate or semi-automate annotation is necessary. Our proposed method is used to identify the number of categories in each scene in a semi-automated way.

Based on the obtained results in Phase I and II of the project we recommend in order to easily annotate large datasets to group the scenes based on one of the following criterion: considering the geographical location of the scenes or the architectural characteristics of the urban areas.

7. REFERENCES

[1] Center for Satellite Based Crisis Information DLR-ZKI Available: http://www.zki.dlr.de/rapid_mapping.
[2] A. Popescu, T. Tsikrika, and J. Kludas, "Overview of the Wikipedia Retrieval Task at ImageCLEF 2010". Available:http://clef2010.org/resources/proceedings/clef2010labs_submission_124.pdf.
[3] M. Everingham, L.Van~Gool, C.K.I. Williams, J. Winn, and A. Zisserman, "The PASCAL Visual Object Classes (VOC) Challenge", International Journal of Computer Vision, vol. 88, no. 2, pp.303-338, 2010.
[4] A. Torralba, B.C. Russell, J. Yuen, "LabelMe: Online Image Annotation and Applications", in Proc. of IEEE, vol. 98, no. 8, pp. 1467-1484, August 2010.
[5] CORINE Land Cover – categories. Available: http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2000-clc2000-100-m-version-9-2007/corine-land-cover-2000-classes-and-rgb-color-codes/clc2000legend.xls
[6] Urban Atlas Available: http://www.eurisy.org/documents _Eurisy/20100531_Ljubljana_Presentations/20100601_7_3_Poelm an.pdf.
[7] C.O Dumitru, J. Singh, and M. Datcu, "Selection of Relevant Features and TerraSAR-X Products for Classification of High Resolution SAR Images", in Proc. of The 9th European Conference on Synthetic Aperture Radar (EUSAR), Nuremberg, pp. 243-246, April 2012.
[8] C.O Dumitru and M. Datcu, "Diversity of Settlement Categories in Very High Resolution SAR Images", in Proc. of Joint Urban Remote Sensing Event 2013, Brazil, pp. 87-90, April 2013.
[9] Virtual Observatory Infrastructure for Earth Observation Data - TELEIOS project, Available: http://www.earthobservatory.eu /knowledge-discovery-from-EO-data
[10] C.O Dumitru and M. Datcu, "Information Content of Very High Resolution SAR Images: Study of Feature Extraction and Imaging Parameters", IEEE Transactions on Geoscience and Remote Sensing, *"to be published"*.
[11] C.O Dumitru and M. Datcu, "Study and Assessment of Selected Primitive Features Behaviour for SAR Image Description", in Proc. of IGARSS 2012, Munich, pp. 3596-3599, July 2012.
[12] C.O Dumitru and M. Datcu, "Information Content of Very High Resolution SAR Images: Study of Dependency of SAR Image Structure Descriptors with Incidence Angle", International Journal on Advances in Telecommunications, vol. 5, no. 3 & 4, pp. 239-251, 2012.
[13] TerraSAR-X Basic Products Specification Document, Issue: 1.6, TX-GS-DD-3302.
[14] MPEG 7, 2012. Available: http://mpeg.chiariglione.org/ standards /mpeg-7/mpeg-7.htm.
[15] Google Maps, 2012, Venice-Italy, http://goo.gl/maps/aoECt.